



Publicació de dades de recerca en obert



Formadora: Cristina Azorín
Moderadora: Marta Jordán

(Maig 2021)

UAB Universitat Autònoma
de Barcelona



Servei de
Biblioteques UAB

Normes i recomanacions

- Teniu els micros silenciats per a una millor comprensió de les explicacions de la formadora
- Teniu prohibit de gravar (vídeo o àudio), directament o indirectament, totalment o parcialment, la sessió.
- Podeu deshabilitar la vostra càmera si així ho desitgeu
- Feu les preguntes a través del xat. S'aniran contestant al llarg de la sessió
- Sentireu la veu de la formadora mentre visualitzeu la presentació

De què parlarem?

DADES DE RECERCA

01

QUÈ SÓN I
TIPOLOGIA

02

OBLIGATORIETAT

03

PRINCIPIS
FAIR

04

EL PLA DE
GESTIÓ DE
DADES

05

EL DDD

Dipòsit Digital
de Documents
de la UAB

06

PUBLICAR
DADES AL
DDD

• Requeriments

07

ALTRES
REPOSITORIS
Per a les dades
de recerca

08

BONES
PRÀCTIQUES
En la publicació
de les dades



Ciència Oberta: La investigació i les dades de recerca accessibles i obertes a tots els ciutadans

Open Science

Open Repositories
Repositoris oberts

Open Access Journals
Revistes d'accés obert



Open Access
Accés obert

Accés sense barreres econòmiques, tecnològiques o jurídiques a les publicacions científiques



Open Peer Review
Revisió per parells oberta



Open Science Evaluation
Avaluació de la ciència en obert

Avaluació oberta dels resultats d'investigació, ampliant la revisió tradicional amb la contribució de la comunitat



Open Metrics and Impact
Impacte i mètriques obertes

Open Reproducible Research
Investigació reproducible en obert

Accés lliure als elements experimentals per a la reproducció de la investigació



Open Research Data
Dades d'investigació obertes



Open Source in Open Science
Codi obert per a la ciència oberta

Open Data
Dades obertes

Dades disponibles en línia de forma gratuïta i que es poden fer servir, reutilitzar i distribuir



Open Big Data
Dades massives obertes



Open Government Data
Dades governamentals obertes



El personal de la teva Biblioteca et pot assessorar

Què són les dades de recerca?

Segons la Comissió Europea:

- fan referència a la informació, **factual** o **numèrica**
- recollida per ser **examinada** i **considerada**
- serveix de base per al **raonament**, la **discussió** o el **càlcul**

Exemples: estadístiques, resultats d'experiments, mesures, observacions resultants del treball de camp, resultats d'enquestes, enregistrament d'entrevistes i imatges

Tipologia de dades de recerca

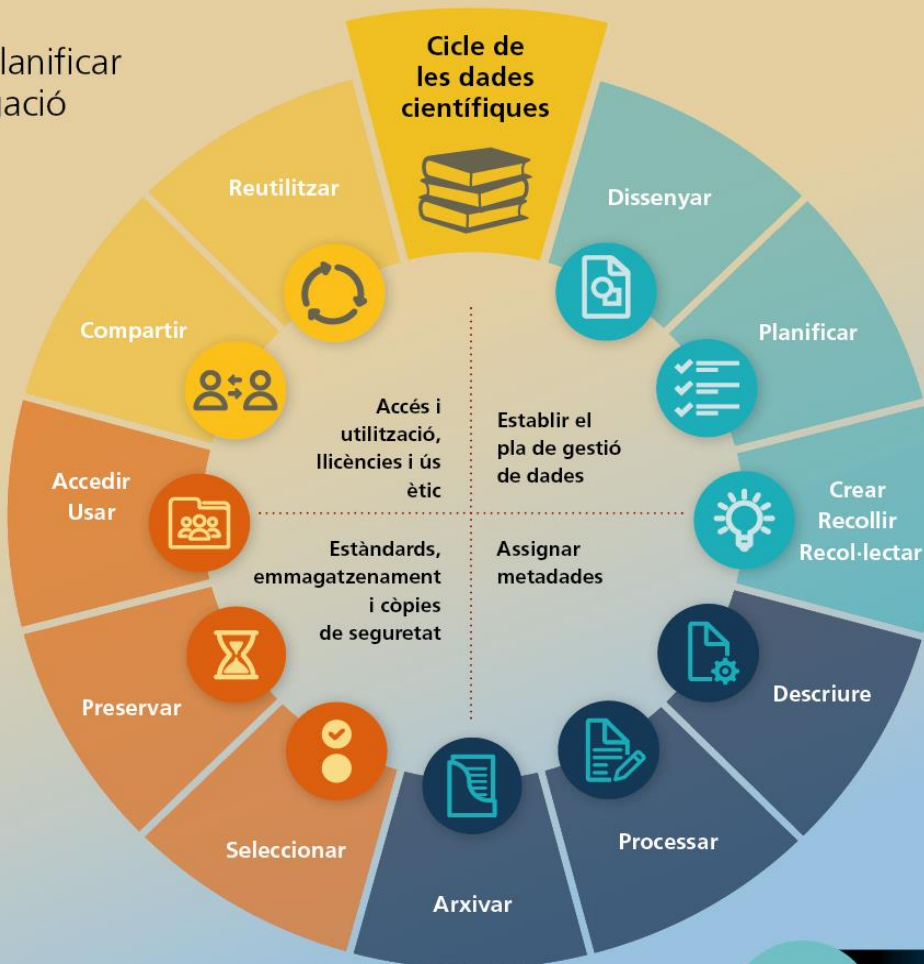
- Dades generades en el procés d'investigació o **dades primàries**: cal ordenar-les, documentar-les i rebre assessorament en el cas de tractar-se de dades personals o sensibles
- Dades associades a resultats publicats o **dades finals**: cal identificar-les, oferir-ne l'accés i facilitar-ne la reutilització



El cicle de les dades científiques



Ajuda a planificar
la investigació



crue

Universidades
Españolas

Red de Bibliotecas
REBIUN



El personal de
la teva Biblioteca
et pot assessorar



Beneficis de publicar les dades

- Reforça la **ciència oberta**: dades disponibles lliurement
- Visibilitat i augment de l'**impacte**
- Identificació amb **DOI** que permet la citació i les mètriques
- Permet la verificació i reproductibilitat dels resultats:
transparència de la recerca
- Promociona la innovació a través de la **compartició** i la **reutilització** de les dades
- Evita la duplictat en l'obtenció i recollida de dades:
optimització de temps, cost i esforços
- Facilita la **col·laboració** i el **debat**: anima la diversitat d'anàlisi i opinions

Convocatòries – Horizon 2020 (2014-2020)



Com complir amb els mandats sobre gestió i publicació de dades

Programa Horitzó 2020 (art. 29.3)



A qui afecta

Investigadors amb projectes subvencionats per l'Horitzó 2020



Per raons de confidencialitat, seguretat, explotació industrial ... **podeu no publicar les dades**



Què obliga a dipositar

Les dades, incloses les metadades, necessàries per validar els resultats presentats en les publicacions científiques

Altres dades, incloses les metadades, especificades als plans de gestió de dades dels projectes d'investigació

Avantatges

- Permetre **validar els resultats** presentats a les publicacions científiques i altres fonts d'informació
- Permeten basar-se en els **resultats d'investigacions prèvies**
- **Fomenta la col·laboració** i evita la duplicació d'esforços
- **Accelera la innovació**
- Millora la **transparència del procés científic**



Requisits

Desenvolupar i mantenir un Pla de Gestió de Dades

Dipositar les dades en un repositori de dades d'investigació

Indicar quines eines cal per fer servir les dades

Permetre l'accés, explotació i disseminació de dades

Es recomana l'ús de llicències Creative Commons



Convocatòries – Horizon Europe (2021-2027)



Open Science across the programme

Open
Science

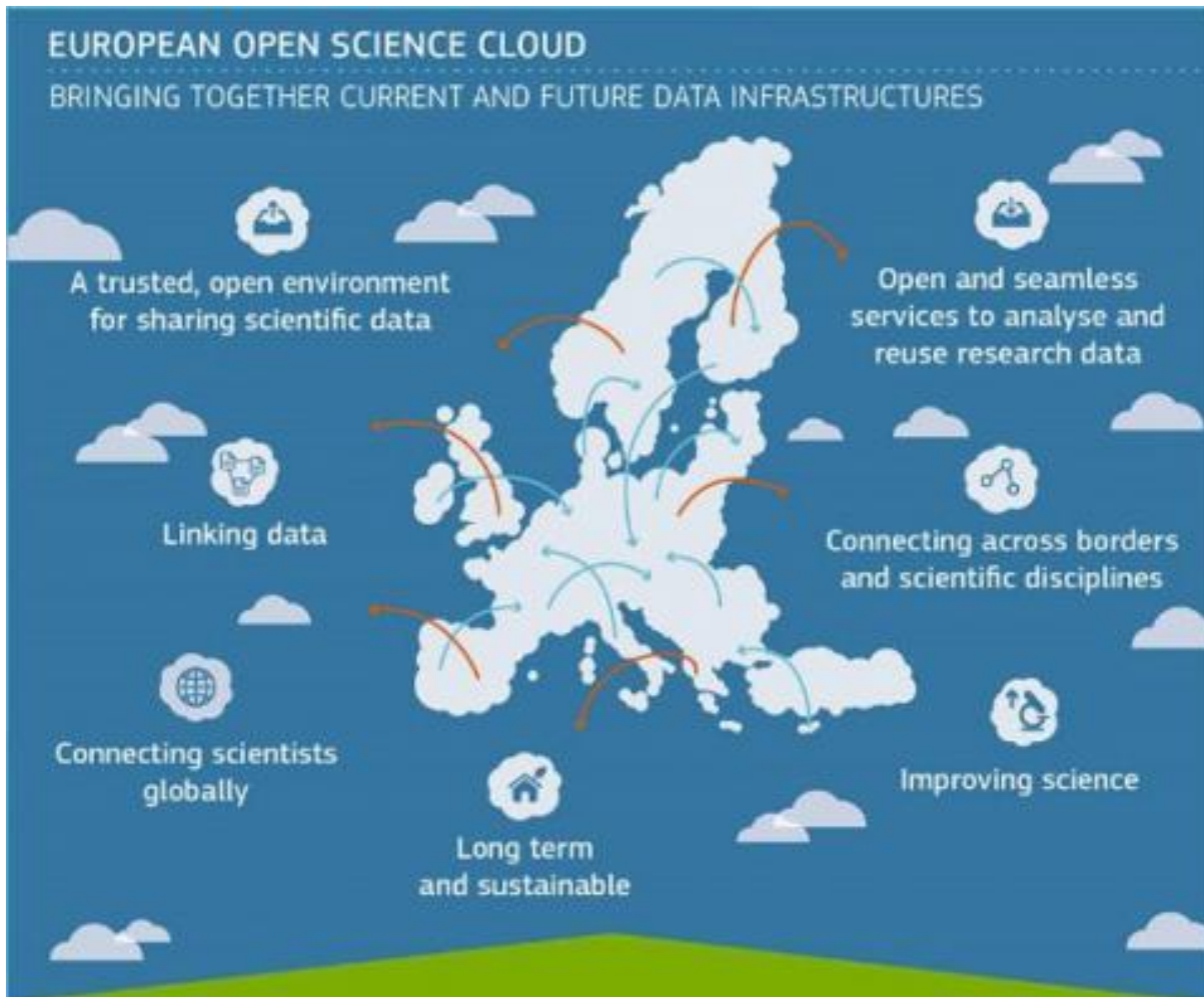
Mainstreaming of open science practices for improved quality and efficiency of R&I, and active engagement of society

Mandatory immediate Open Access to publications: beneficiaries must retain sufficient IPRs to comply with open access requirements;

Data sharing as ‘open as possible, as closed as necessary’: mandatory Data Management Plan for FAIR (Findable, Accessible, Interoperable, Reusable) research data

- Work Programmes may incentivize or oblige to adhere to **open science practices** such as involvement of citizens, or to use the **European Open Science Cloud**
- Assessment of open science practices through the **award criteria** for proposal evaluation
- Dedicated support to **open science policy actions**
- **Open Research Europe** publishing platform

European Open Science Cloud



eosc-portal.eu

Convocatòries – Plan Estatal (2017-2020)

GENERACIÓN DE CONOCIMIENTOS E INFRAESTRUCTURAS DE INVESTIGACIÓN

- Posicionamiento científico internacional (O1)
- Participación en H2020, grandes proyectos e infraestructuras de investigación internacionales (O2)
- Capacidad tractora de centros e instituciones sobre el conjunto del Sistema (O3)
- Valorización de la I+D+i pública (R1)
- Impacto científico y social (R2)
- Fortalecimiento y acceso a ICTs, interoperabilidad y modernización de infraestructuras y equipación (R3)
- Redes de colaboración científica e integración transdisciplinar (R4)
- Difusión de los resultados a la sociedad (R5)
- Acceso Abierto y Ciencia Abierta (R5)



Els projectes poden incloure, *optativament*, un Pla de Gestió de dades de recerca i el dipòsit de les dades en repositoris institucionals

Estrategia Española de Ciencia, Tecnología e Innovación (2021-2027)

CIENCIA E INNOVACIÓN EN LA SOCIEDAD

Eje 14. **Social.** Promover el compromiso de la **sociedad española con la I+D+I**, fomentando la **divulgación y la cultura científica**, la reflexión sobre el papel de la ciencia y la tecnología en la sociedad actual, y promoviendo una ciencia e innovación **abiertas e inclusivas**.



El apoyo a la **ciencia de excelencia**, elemento básico del modelo de EECTI, fomentará, en consonancia con las directrices de la UE, el **acceso abierto a los resultados de investigación**, permitiendo que los **datos sean accesibles, interoperables y reutilizables** (su acrónimo en inglés *FAIR*). La difusión en el ámbito científico, junto al esfuerzo llevado a cabo por los **repositorios abiertos**, facilitará la accesibilidad a los avances científicos y fomentará la divulgación y comunicación científica hacia la sociedad, objetivo que se persigue en el **Eje de actuación 14**.

Política institucional específica



- **Preàmbul:** es reconeix la importància de les dades de recerca a la UAB
- **Jurisdicció:** és aplicable a tot el personal d'investigador en actiu
- Aspectes relatius al **dipòsit de les dades:** repositori, preservació, compromís FAIR, període de preservació...
- Cal establir **una llicència** que permeti una reutilització àmplia, com per exemple, les llicències CC0 o CC BY de Creative Commons
- **Responsabilitats** dels agents implicats: els investigadors i la institució
- **Validesa:** es revisarà cada 5 anys (aprovada el març de 2020)

ddd.uab.cat/record/222172

Políticas editorials

BMJ Journals

WILEY



SPRINGER NATURE



ELSEVIER

OXFORD
ACADEMIC

Principis FAIR

Findable:

- ✓ Identificador persistent: DOI (preferentment) i/o URI
- ✓ Publicar en repositoris institucionals o temàtics que descriguin les dades emprant estàndards internacionals
- ✓ Els repositoris han de permetre que tots els seus continguts siguin cercables i recuperables



Principis FAIR

Accessible:



- ✓ Les dades i les metadades seran accessibles perquè els repositoris utilitzen protocols de comunicació estàndards que permeten l'intercanvi d'informació entre plataformes

Principis FAIR

Interoperable:

- ✓ Les metadades han de complir uns estàndards internacionals (Dublin Core, DataCite...), tant pel que fa a codis com a nomenclatures.



Principis FAIR

Reusable:

- ✓ Les dades han d'estar descrites exhaustivament i en detall: origen, data de recollida, circumstàncies de la recollida, observacions...
- ✓ Les dades han d'estar descrites preferentment en estàndards de l'àrea temàtica corresponent. Si això no és possible cal utilitzar un estàndard general àmpliament reconegut
- ✓ Hi ha d'haver una menció de drets per a cada dataset (licències CC o en domini públic)



Pla de Gestió de Dades (PGD)

El Data Management Plan (DMP) descriu el cicle de vida de la gestió de les dades que es generen durant el procés de recerca

El pla hauria d'incloure informació sobre:

- la gestió de les dades durant i després del final del projecte
- quines dades es generaran, recopilaran i processaran
- quina metodologia i estàndards s'aplicarà en la gestió
- quines dades estaran disponibles en accés obert
- com es preservaran les dades durant i després del final del projecte



10 passos per elaborar un Pla de Gestió de Dades

Un **Pla de Gestió de Dades** (PGD) o Data Management Plan (DMP) és un **document formal**, que es presenta a l'inici de la investigació, en el qual es

descriu què es farà amb les dades al llarg i després de finalitzar la investigació i que es pot modificar si es produeixen canvis en la mateixa.

Per què?

És una **bona pràctica**, és un **element clau de l'Open Science** i és **obligatori** als nous projectes H2020.

Eines gratuïtes per elaborar un PGD



PGDonline
(Consorcio Madroño)
<http://dmp.consorciojadrono.es/>



DMPOnline (Digital Curation
Centre, UK)
<https://dmponline.dcc.ac.uk/>



Reviseu els **requeriments** de l'ens finançador (H2020).



Identifiqueu les dades: tipologia, procedència, volum, formats i fitxers.



Definiu com s'organitzaran i gestionaran les dades: nom dels fitxers, control de versions, programari necessari...



Expliqueu com es documentaran les dades: identifiqueu la informació a processar, consulteu si hi ha estàndards o esquemes de metadades, identifiqueu eines per a gestionar-los.



Descriu els processos que asseguruen una **bona qualitat de les dades**.



Prepareu una estratègia d'emmagatzemament (durant el procés) i de preservació de dades (repositori).



Definiu les polítiques de dades del projecte: qüestions sobre propietat intel·lectual i com es tractaran les dades sensibles i personals.



Descriu com es difondran les dades: on, quines, quan es difondran. Si publicareu les dades en un repositori, com a informació suplementària de l'article o com un "data paper".



Assigneu rols i responsabilitats a les persones i organitzacions participants en el projecte.



Prepareu un pressupost realista: la gestió de dades costa temps i diners en termes de programari, maquinari, serveis i personal



crue Universidades Españolas Red de Bibliotecas REBIUN

Adaptat de Michener, WK. Ten simple rules for creating a good data management plan. PLOS Comp Biol. 2015; 11(10)



El personal de la teva Biblioteca et pot assessorar



PGD per a la tesi doctoral

Consorci de Serveis Universitaris de Catalunya (CSUC)

Guia per elaborar un pla de gestió de dades per a doctorands

ddd.uab.cat/record/218680

Eina DMP – plantilla que facilita l'elaboració del PGD

dmp.csuc.cat

eiNa DMP
Pla de Gestió de Dades de Recerca

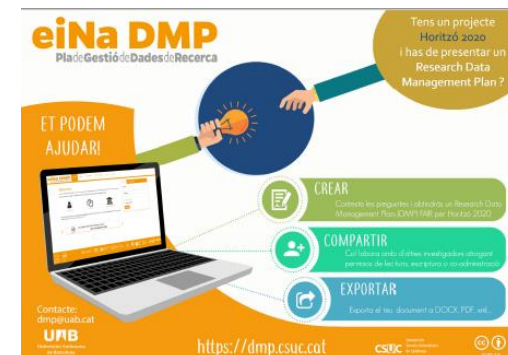
PGD – Material de suport



- Videotutorial CSUC: youtu.be/8gKSePSVkpl
- Videotutorial de la UPF: youtu.be/s5JF7LPCMjw



- Infografia: ddd.uab.cat/record/181194



De què estem parlant?



Repositoris

*Un dipòsit digital de confiança és aquell que té com a missió proporcionar **accés fiable** i a **llarg termini** als recursos digitals gestionats a la comunitat designada, ara i en el futur.*

El Dipòsit Digital de Documents UAB



Datasets

- ddd.uab.cat/collection/datasets

Plans de Gestió de Dades

- ddd.uab.cat/collection/plagesdad

El DDD és...

| | | |
|---|--|---|
| F | Assigna identificadors (DOI/URI) | ✓ |
| | Empra estàndards internacionals | ✓ |
| | Continguts cercables i recuperables | ✓ |
| A | Dades i metadades accessibles a través de protocols de comunicació estàndard | ✓ |
| I | Metadades estàndard | ✓ |
| R | Descripció exhaustiva de les dades | ✓ |
| | Menció de drets (licències CC o domini públic) | ✓ |
| | Descripció segons estàndards temàtics o generals àmpliament reconeguts | ✓ |

Publicar dades al DDD

- ✓ identificador ORCID
- ✓ resum del contingut (pot ser en diversos idiomes)
- ✓ paraules clau (poden estar en més d'un idioma)
- ✓ marcar la llicència d'ús ([recomanacions de la UAB](#))
- ✓ indicar el codi del projecte (en cas de gaudir d'un ajut per a la recerca)
- ✓ indicar l'obra relacionada (article, tesi, llibre...)
- ✓ 25 Gb cada dataset (si necessiteu més espai, contacteu amb la vostra biblioteca)



Altres repositoris

A banda del DDD, hi ha d'altres repositoris que permeten hostatjar datasets



Trobareu les característiques de cada repositori a:
ddd.uab.cat/record/150829

Altres repositoris – criteris per a la tria

1. Existeix algun repositori representatiu i consolidat en el teu àmbit?
2. El repositori permet depositar les dades en el format utilitzat?
3. La mida dels fitxers que pots depositar s'ajusta a les teves necessitats?
4. Es permet la restricció d'accés a les dades (tancades, restringides o embargades)?
5. Els drets i llicències que t'ofereix el repositori s'ajusten amb els teus interessos?
6. El repositori atorga un identificador permanent?
7. Les condicions per depositar o retirar contingut del repositori s'ajusten a les teves necessitats?
8. Hi ha costos associats a l'ús del repositori?
9. Es poden enllaçar les dades a les publicacions associades?
10. El repositori té una política de preservació (còpies, terminis de conservació)?
11. Permet depositar diferents versions d'un mateix arxiu?
12. Compleix els requisits dels ajuts europeus?

Bones pràctiques



- ✓ afegir un arxiu tipus README on detalleu el dataset (ddd.uab.cat/record/226740)
- ✓ noms de fitxers identificables (ddd.uab.cat/record/233037)
- ✓ arxius en formats oberts (txt, csv, pdf, LaTeX, jpg, MP3...)
- ✓ un PGD que acompanyi les dades
- ✓ citeu les dades, les vostres i les que utilitzeu de tercers (ddd.uab.cat/record/233039)
- ✓ cada conjunt de dades ha de tenir un únic DOI

Recordeu!



**Qualsevol persona que consulti les vostres dades,
ha d'entendre-les per poder-les reutilitzar**

Exemples

- ✓ Relació entre el document i les dades al DDD

ddd.uab.cat/record/232224

Obra relacionada: Mendoza Domínguez, Núria. La formación en audiodescripción en España
<http://hdl.handle.net/10803/671267>

DOI: [10.5565/ddd.uab.cat/232224](https://doi.org/10.5565/ddd.uab.cat/232224)



Audiodescriptor
Fichero xlsx
43.5 KB



Prácticas
Fichero xlsx
42.4 KB



Transcripciones



README
Fichero txt
1903x996, 48.5 KB

- ✓ Disponibilitat de les dades al document publicat

<https://doi.org/10.1007%2Fs10021-020-00500-z>

Data Availability

Data supporting the findings of this study (DBD, water content, grains size distribution, C and N contents, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$, ^{210}Pb and DOC) are available at <http://ddd.uab.cat/record/216456> with the identifier <https://doi.org/10.5565/ddd.uab.cat/216456>



Cita les teves dades d'investigació



Per què és important citar les dades:

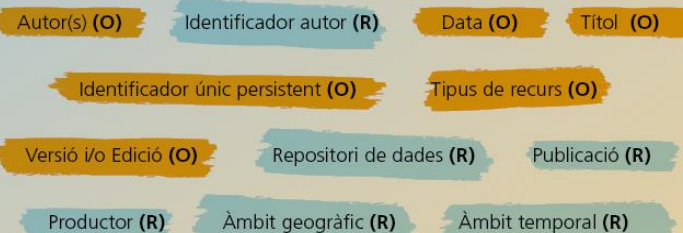
- Els conjunts de dades també són resultats de la recerca com els articles, monografies, etc.
- Facilita la identificació i l'accés a les dades i així facilitarà la seva localització, validació i reutilització
- Permet conèixer l'autoria dels seus creadors
- Facilita la mètrica i impacte de les dades
- Afavoreix la transparència de la investigació científica

Bones pràctiques per citar dades:

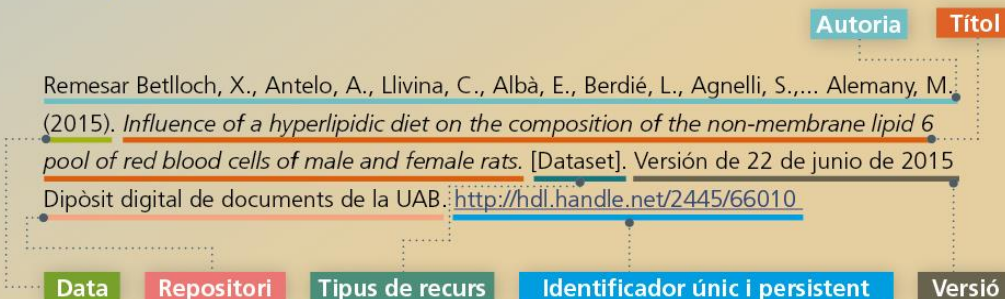
- Cal facilitar la identificació, localització i l'accés a les dades mitjançant un identificador únic i persistent (DOI, Handle, etc.).
- Cada conjunt i subconjunt de dades (dataset) s'ha de citar de forma independent.
- Les cites de les dades usades han d'aparèixer a la secció de referències bibliogràfiques de la publicació resultant.
- Es recomana incloure un identificador únic d'autor (ORCID, etc.).

Elaboració de la cita

- Hi ha elements mínims obligatoris (O) i altres recomanats (R) que es combinen per elaborar la cita en qualsevol estil estàndard (APA, MLA, Chicago, etc.) o bé els proposats pels principals repositoris de dades (Dataverse, Dryad, etc.).



Exemple de cita estil APA



Enllaça les dades amb els documents resultat de la investigació i a l'inrevés, i crea les referències bibliogràfiques dels mateixos.



crue

Universidades Españolas

Red de Bibliotecas REBIUN



El personal de la teva Biblioteca et pot assessorar



DADES DE RECERCA A LA UAB

Open Access UAB



OPEN ACCESS

Universitat Autònoma de Barcelona > Investigar > Open access UAB > **Producció científica en obert**

Accés obert

Presentació

Producció científica en obert

Termòmetre de la recerca en obert a la UAB

Amb qui contactar

Comissió de l'Accés Obert

Com dipositar en obert

Dades de recerca en accés obert

Drets d'autor

Horizon 2020

Producció científica en obert

Activitat del DDD



Lliura



Intranet - PDI



www.uab.cat/open-access

Bibliografia

- Consorci de Serveis Universitaris de Catalunya. Recomanacions per seleccionar un repositori per al dipòsit de dades de recerca [Internet]. 2019 [Consulta 3 abril 2021]. Disponible a: <https://ddd.uab.cat/record/163562>
- España, Ministerio de Ciencia e Innovación. EECTI: Estrategia Española de Ciencia, Tecnología e Innovación 2021-2027 [Internet]. [Madrid]: Secretaría General Técnica del Ministerio de Ciencia e Innovación, [2020] [Consulta 1 abril 2021]. Disponible a: <https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/EECTI-2021-2027.pdf>
- España, Ministerio de Economía, Industria y Competitividad. Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020 [Internet]. [Madrid]: Gobierno de España, [2017] [Consulta 15 gener 2020]. Disponible a: <http://www.ciencia.gob.es/stfls/MICINN/Prensa/FICHEROS/2018/PlanEstatalIDI.pdf>
- European Commission. H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020 [Internet]. [Brussels]: European Comission, 2016 [Consulta 3 Abril 2021]. Disponible a: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

- European Commission. H2020 Programme: Guidelines to the rules on Open Access to scientific publications and Open Access to research data in Horizon 2020 [Internet]. [Brussels]: European Commission, 2017 [Consulta 1 Abril 2021]. Disponible a: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- European Commission. Horizon Europe: the EU Research & Innovation programme 2021-27 [Internet]. [Brussels]: European Commission, 2021 [Consulta 3 Abril 2021]. Disponible a: https://ec.europa.eu/info/sites/info/files/research_and_innovation/funding/presentations/ec_rtd_he-investing-to-shape-our-future.pdf
- European Commission. European Open Science Cloud [Internet]. [Brussels]: European Commission, 2021b [Consulta 3 abril 2021]. Disponible a: <https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud>
- LEARN project, ed. LEARN Toolkit of best practice for Research data Management [Internet]. [s.l.]: Leaders Activating Research Networks, 2017 [Consulta 3 Abril 2021]. Disponible a: <http://dx.doi.org/10.14324/000.learn.00>
- Red de Bibliotecas Universitarias Españolas (REBIUN). Datos de investigación y acceso abierto [Internet]. [Consulta 4 Abril 2021]. Disponible a: <https://ddd.uab.cat/record/165475>
- Zenodo. FAIR principles [Internet]. Genève: CERN, [201?] [Consulta 4 Abril 2021]. Disponible a: <http://about.zenodo.org/principles/>

Podeu consultar aquesta presentació a:
<http://ddd.uab.cat/record/222106>

Per acabar

- Rebreu per correu-e:
 - Enllaç a la **presentació** per consultar-la i descarregar-la.
 - Enllaç a l'**enquesta** per ajudar-nos a millorar
 - El **certificat** d'assistència
 - Preguntes i respostes fetes al xat
- Assessorament personalitzat a: **Pregunt@**
<https://www.bib.uab.cat/pregunta/pregunta.php>



Gràcies!

#bibliotequesUAB

